DOCUMENT RESUME

ED 325 523 TM 015 781

AUTHOR Tittle, Carol Kehr
TITLE Technology and Accountability: Does More Equal Better? Should Accountability Programs Be Accountable?
PUB DATE 1 Nov 90
NOTE 12p.; Revised version of a paper presented at the Annual Urban Education Technology Forum Conference of the Council of the Great City Schools (5th, New York, NY, October 17, 1990).
PUB TYPE Viewpoints (120) -- Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Accountability; Educational Assessment; Educational Change; Educational Improvement; *Educational Technology; Elementary Secondary Education; Evaluation Methods; Higher Education; Measurement Techniques; Program Evaluation; Research Universities; *School Effectiveness; Standardized Tests
IDENTIFIERS National Assessment of Educational Progress

ABSTRACT

The role of technology and the measurement of educational progress are discussed, with an emphasis on the development of school report cards. Comments are made from the perspectives of (1) a member of a university community and (2) a researcher in evaluation and measurement. Efforts to formulate school report cards in Florida, California, and elsewhere make it apparent that universities can look forward to extensions of recordkeeping and reporting for accountability purposes. An evaluation perspective leads to two concerns: (1) the view of the organizational change or school improvement that is assumed by the school report card; and (2) the need for reflection on and evaluation of the school report card programs. Standardized tests have constituted a key element in school report cards. The National Assessment of Educational Progress is an example of the direction that testing should take. Developing assessments that are meaningful to teachers and learners will require long-term cooperative efforts of teachers and specialists. It is concluded that more information does not necessarily mean better schools or postsecondary education. Technology does not determine the choices for what a school or postsecondary report card contains or the system/legislation that fosters the report card. The challenge is to study the meaning and use that is derived from accountability measures; develop alternative or multiple assessments, and change from thinking that a single accountability system or any accountability system can meet the goal of accountability and the goal of school improvement. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED325523

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
CAROL KEHR TITTLE

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

c21cgcsr

November 1, 1990

Technology and Accountability: Does more equal better? Should accountability programs be accountable?

Carol Kehr Tittle
Graduate School
City University of New York

Revised version of a Paper presented at a panel on School Report Cards and the Role of Technology, Council of the Great City Schools' Fifth Annual Urban Education Technology Forum Conference, October 17, 1990, New York City.

I have been asked to comment on the role of technology and the measurement of progress, particularly on the development of school report cards. These comments will be made from two perspectives: the first is the perspective of a member of the university community; and the second is the perspective of a researcher in evaluation and measurement. From the first perspective, schools and universities face a common legislative policy of accountability. From the second perspective, the view of organizational change and improvement guiding the current school report card policy and legislation is one that is top-down, information-driven and assumes a rational process of change. Both perspectives lead to a recommendation for reflection on accountablity--for evaluation of the accountability systems.

Background

By way of background, school (and university) report cards are the latest push in the accountability movement that has been underway since the federal initiative in education in the 1960's and 1970's. Much of the federal expansion or increased funding role in the 1960s was accompanied by evaluation requirements that became more and more prescriptive. Evaluation data were intended to provide the basis for policy decisions, to retain, expand, or curtail programmatic federal efforts in education.

The federal initiative expanded state departments of education and increased both the number of professional staffs and their levels of expertise. Further expansion of state level staffs occurred as a function of the increased proportion of local education funds provided by state revenues in the 1970's and 1980's. The shift in funding and staffing at the state level has been accompanied by more accountability programs designed at the state level, from minimum competency and basic skills state-wide testing

TM015781

programs to the ambitious school outcome and demographic data compilation efforts of the school report card.

While there has been an expansion of accountability programs, a critical tension has arisen. This tension results from the collision of curriculum reform efforts and the drastically expanded use of standardized, multiple choice tests that accompanied the increase in accountability programs. The emerging "revolution" in assessment now underway in the curriculum areas, especially mathematics, science, reading, and writing, is an effort to change the form of assessment in accountability programs and thereby influence practice. The effort to change the approach to assessment attests to the power of the accountability programs to influence practice at some level. This change in assessment procedures does not appear to be a challenge to changes in the basic characteristics of the accountability programs: every pupil testing, with class, school, district, and state level aggregations of data.

The speed of the change in language about assessment has been startling: states and LEAs are struggling to define "authentic" assessment and performance assessment in the curriculum areas. State-wide efforts in performance assessment are evident in Connecticut, California, and Maryland. Clearly, in the measurement or assessment of outcomes, if you will, earnest efforts toward change are in process.

How does the development of accountability programs and even changes in assessment procedures relate to technology and school report cards? The report cards are part of an on-going movement. The technology has provided an underpinning that supports more powerful applications of data collection, analysis and use, and technology has potential in assessment itself. However, the tension with the curriculum reform efforts indicates that there is more at stake than changing the form of assessment in state or citywide accountability programs. The curriculum reforms are part of a broader and fundamental change in our views about learners and teachers. Business-as-usual in the accountability programs needs examination.

The challenge, as I see it, for the members of the Council of the Great City Schools is to critically review current and proposed applications of technology in accountability for their policy implications. This review or reflection needs to be conducted with an understanding of the fundamental change in the view of the teacher and learner that is at the heart of the current curriculum reform efforts. We need to ask several questions: Does more equal better? And more particularly, can we define what we mean by better? How wise is a policy that focuses efforts for school improvement on expanded accountability measures as

reflected in the formulation of school report cards? What alternatives can we create?

The university perspective.

Let me comment briefly on how universities are drawn into the school report card discussion. They are drawn in by two developments: 1) the requirements that school report cards include in some cases data on students as they move to postsecondary education; and 2) the requirements of the U.S. Congress for a report card for postsecondary institutions. In the first development, several states have been at the forefront of the effort to define school report cards. Examples are the legislation in Florida, Tennessee, and California.

In Florida, institutions of postsecondary education are required to report to the Commissioner of Education each term, information on first-time-in-college Florida high school graduates. This information includes performance on regular courses and referrals for remediation in postsecondary education. In Tennessee (McLarty & Hudson, 1987), there is a statewide commissioner's report card and optional school report cards are apparently also in use (The Tennessee School Improvement Project, TSIP). The commissioner's report card includes student outcome data, test scores and diplomas; student process data, AP courses, attendance rates, suspensions, etc.; and other data such as teacher characteristics, licenses; teacher/classroom process data, teacher/pupil ratios.

The optional TSIP is more ambitious than the commissioner's state-wide report card and uses the framework from the more effective schools research literature: a strong instructional focus, a safe and positive school climate, frequent monitoring of student progress, a strong instructional leader, high expectations for student achievement, and strong parent and community involvement), and is optional for schools. The TSIP is viewed as formative in nature--providing information for guidance of the improvement process. (A paper related to the Tennessee project is by Furtwengler, 1986.)

California has an elaborate system (Fetler, 1989). The California report card provides for three types of comparisons: a school with itself, over time; with all schools statewide; and with schools with similar demographic characteristics (based on a composite SES indicator using percent AFDC, parent education or occupation; percent LEP, student mobility). In the last case, a school is compared with the 10% above it and the 10% below it in ranking on the composite SES indicator). It is important to note that California proposes to directly examine equity issues.

Equity issues are addressed by presentations of data by sex and by major ethnic groups.

Variables on the California report card can be grouped into three areas:

1. Status characteristics such as those found in traditional accreditation reports--teachers in license, class size, per pupil expenditures, teacher training, etc.; average salaries of teachers, principals, and superintendents.

2. Outcome variables--achievement, California Assessment Program, locally administered measures, SAT, ACT, drop out rates; enrollments at postsecondary institutions.

3. Process variables--quality of textbooks, classroom discipline and climate, quality of instruction, quality of leadership.

These are a combination of types of data, as with the TSIP. At risk schools are identified and there is a program and time line for improvement.

Fetler (1990) describes the political process in California by which the report card information was negotiated with the Association of California School Administrators and the State Department of Education. He raised the issue of "...whether accountability is best seen as a rational technique for managing change in schools or is it primarily a political symbol" (1990, p.11). He provides examples of where each group--parents, administrators, teachers, legislators, etc., represented their interests, and concludes, "...Even if this result is not the most rational or effective that could be imagined, presumably it still has value in getting people together to work for positive ends" (1990, p. 11).

To summarize, because of the follow-up data on high school graduates required in school report cards, universities in Florida, California, and elsewhere can undoubtedly look forward to extensions of record keeping and reporting for accountability purposes.

The second development drawing universities into the record card accountability program is more direct. Congress is expected to pass a measure requiring colleges and universities nationally to tell prospective students and their parents the graduation rates and crime statistics on their campuses (New York Times, Monday, October 8, 1990, pA1,A15). Congressional negotiators have reached agreement on reporting crime statistics and graduation rates for

athletes by area of sport. The Education Department will be directed to study the usefulness and feasibility of compiling figures on graduation rates broken down by academic discipline, job placement, and job certificate rates for students, and revenues and spending in athletic programs (Chronicle of Higher Education, Congressional Negotiators Approve Graduation-Rates Bill, October 17, 1990, P. A23, A28). The first reports are due September 1, 1992.

This bill evolved from an original measure that required only that colleges disclose the graduation rates of their athletes. The reasoning was stated by Senator Edward M. Kennedy: "Sunlight is the best disinfectant. Once colleges begin disclosing this vital information, those with the poorest records will be under the greatest pressure to improve" (New York Times, 10/8/90,p.A1).

As with the Florida legislation requiring rankings of schools on a nationally normed testing program, the Congress expects that newspapers and others will publish the data comparing colleges. The data are viewed also as consumer information, and probably with more justification for postsecondary institutions than for schools. Problems have been noted in defining what is a campus, for purposes of the crime statistics, and what graduation rates mean for institutions with different populations. Such problems of interpretation and definitions will sound familiar to school administrators who have been confronted with drop out rate statistics and where comparisons are made for schools with students who have different mobility patterns. (The Florida legislation does provide for reporting transient students separately for drop out rate purposes for districts and schools, 1990 Legislature, CS/NB 931, Commissioner's Report, pages 139-145.)

While expanded technology supports the gathering of such data, it will require extensive effort on the part of all concerned to ensure comparability of records across institutions and the preparation of accurate large scale data bases. Further, the preparation of information such that accurate inferences and meaning are derived from their publication needs study. This concern with inferences and meaning leads to discussion below of evaluating the accountability programs with school record cards.

The evaluation perspective.

I want to mention two concerns from an evaluation perspective. The first is the view of organizational change or school improvement that is assumed by the school report card policy. The second is the need for reflection on and evaluation of the school report card programs, including a

validity-inquiry framework to guide that reflection and evaluation.

First, What is the view of organizational change on which school report cards are based? Organizational change is assumed to be a rational process, responsive to information of a particular sort: The information that can be collected, analyzed and reported by counting and marking. This view also assumes that the school report card data have meaning to all users, that is, that users will make accurate inferences. It also assumes that the school report card data will be appropriately used. These views can be characterized as technological in orientation--the collection of data is good in and of itself, without understanding the context of professional practice. To my knowledge, the assumptions of meaning and use are not substantiated by research. Research on the meaning teachers and students derive from standardized test results does not suggest that the assumptions are warranted (e.g., Tittle, Kelly-Benjamin & Sacks, in press).

To a considerable degree, it is also a view of organizational change as imposed from the top-down, in terms of the requirements for several of the school report cards. An exception is the Tennessee formative evaluation effort, with the Tennessee School Improvement Project. Overall, the policy view still seems to be what McLaughlin (1989) described as implementation, rather than enabling. In McLaughlin's view, policy makers do not acknowledge the need to move beyond formal organizational policy and structures to promote improvement and stimulate change.

The federal efforts in evaluation have, over the long term, yielded some implications for policy aimed at school improvements. As McLaughlin summarizes these implications, they are:

special projects or reforms aimed at discrete elements of the education policy system are likely to disappoint;

policy intending to promote more effective educational practice must address both (content & process), and acknowledge the need for the quite different kind of expertise associated with the management of organizational change and with improved content;

there is the need to look beyond formal policy structures for channels for promoting improvement and stimulating change--e.g., urban math collaboratives; and

a focus on removing constraints (to better practice) is not the same as enabling practice, which may require different factors such as productive collegial relations and

organizational structures that promote open communication and feedback.

In McLaughlin's view, the focus is moving from understanding policy implementation to enabling effective practice. This focus underscores the essential contribution of teachers' perspectives as informants and guides to policy.

The insights resulting from reflecting on the evaluations over time of federal programs suggest the merit of an effort to reflect on and evalute the expanding accountability programs that technology has made possible. The need for reflection on school record cards and the expanded accountability efforts is evident from research and evaluation that has been conducted on the basic skills and competency testing programs.

Some of the research has examined effects on curriculum and teaching. The state-wide testing programs have indeed driven the curriculum and teaching, but in many instances in a negative direction. There are examples of various degrees of test preparation, coaching, or teaching directly to the standardized tests in use in these programs (see Ellwein, Glass, and Smith, 1988; Mehrens and Kaminsky, 1989; and Shepard, 1990). Teachers in New York city spend months preparing students for the city-wide tests; comments by Fetler (1990) suggest the same problems exist in California. I do not think these problems will go away when "authentic" assessments are used. The arguments that it is better for authentic assessments to drive curriculum and teaching seems to me to be fundementally in conflict with the philosphy of the curriculum reforms and the efforts to professionalize teaching. Further, as Campbell (1977) noted years ago, social indicators that are used for decision making are likely to be corrupted.

The problems with large scale every-pupil testing for accountability purposes are becoming more evident as studies examine the measures both from a technical view point and from the practices of test preparation that occur in the school. Without a knowledge of the "opportunity to practice" the test as a conditional variable, these test results are largely uninterpretable. The manipulation of who takes the tests also contributes to the interpretation problems. For individuals knowledgeable about schools and testing, little trust can be placed in the large scale accountability program results. Even Cannell's Lake Woebegone charge has been supported to a considerable degree by analyses of nationally normed standardized tests (Linn, Graue, and Sanders, 1990).

Reflection on and evaluation of school report cards.

Reflection on and evaluation of school report cards would examine the school report card as part of a system that includes their development, preparation, distribution, and use. The role of the publication of outcome information as a causal agent in change would be examined, and there would be study of the validity of the demographic, process, and outcome information collected and reported. Examples of the focus of validity-related studies include the meaning and inferences:

. that relevant groups in the education and larger community derive from the school report cards, and

. that relevant groups in the education and larger community derive from the information and technical support or training programs that accompany the school report cards.

These studies would also include descriptions of the use of the school report cards:

Do they provide necessary or sufficient information to the different members of the education and larger community involved--including students, parents, teachers, administrators, school boards?

What actions are taken by these different groups?

What are the intended and unintended outcomes and consequences of these actions?

These questions draw on current thinking about validity and the construction of meaning (Messick, 1990; Tittle, 1990).

Alternatives

Having considered problems with the standardized test scores that constitute a key element in the school record cards, and their negative influence on practice, and listed evaluation questions, what are some alternative paths that can be taken?

The National Assessment of Educational Progress (NAEP) stands as an excellent model. Serious reflection on accountability programs and evaluation of these programs should result in insights that may shift the practices and focus of the school report cards. NAEP is a model from the standpoint that the sampling plan used in assessment reduces the pressures to teach to the test and permits more information to be collected in brief time periods. These charactistics of NAEP can meet national (or state) accountability needs. However, such a system does not, and should not (except in a general sense), provide teachers and

individual learners with information about individual progress and needs.

Fresh thinking is needed about the information that professional teachers and learners in classrooms find meaningful and can use, what such information might be, how it could be developed, by whom, and when. Such re-thinking of assessment for teaching and learning can focus on possibilities that cannot be considered in assessment for accountability: assessments developed by teachers in a school, or by math teachers in collaboratives, for example; portfolios of student writings or projects; assessments reviewed by school management committees, or curriculum committees for the school or district.

These assessment alternatives could focus on enabling practice, in McLaughlin's terms, and would reduce the pressure to take any form of assessment and distort its intentions, as occurs in the large scale accountability programs. To develop assessments that are meaningful to teachers and learners is no trivial exercise, and will require long-term, cooperative efforts of teachers and specialists who can provide technical support through state, district and city curriculum and assessment specialists.

Summary

Let me summarize by repeating my main points:

. More information does not necessarily mean better schools or postsecondary education;

. Technology does not determine the choices for what a school or postsecondary report card contains or the system/legislation that fostered the report card;

. The challenge is to reflect on and study the meaning and use that is derived from accountability measures, including school and postsecondary report cards; and

. The challenge is to expand our thinking, to develop alternative or multiple assessments, to change from thinking that a single accountability system or any accountability system,, national, state, or Great City, can meet the goal of accountability AND the goal of school improvement.

References

Campbell, D.T. (1977). Keeping the data honest in the experimenting society. In H.W. Melton & D.J.H. Watson (eds.). Interdisciplinary dimensions of accounting for social goals and social organizations. Columbus, Ohio: Grid.

Ellwein, M.C., Glass, G.V., & Smith, M.L. (1988). Standards of competence: Propositions in the nature of testing reforms. Educational Researcher, 17, 4-9.

Fetler, M. (1989, March). Assessing educational performance: California's school quality indicator system. paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. ED 312 300

Fetler, M.E. (1990, April). Refining a performance accountability system in California: A case study of a constitutional initiative. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.

Furtwengler, W.J. (1986, April). Reaching success through involvement--implementation strategy for creating and maintaining effective schools. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Linn, R.L., Graue, M.E., & Sanders, N.M. (1990). Comparing state and district results to national norms: The validity of the claim that "Everyone is above average." Educational Measurement: Issues & Practice, 9, 5-14.

McLarty, J.R., & Hudson, S. (April 1987). Formative and summative approaches to effectiveness indicators. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

McLaughlin, M.W. (1989, September). The Rand change agent study ten years later: Macro perspectives and micro realities. CRC No. P89-108. Stanford, CA: Center for Research on the Context of Secondary Teaching, School of Education.

Mehrens, W.A., & Kaminsky, J. (1989). Methods for improving standardized test scores: Fruitful, fruitless, or fraudulent? Educational Measurement: Issues & Practice, 8, 14-22.

Messick, S. (1990). Validity. In R.L.Linn (ed.), Educational Measurement (3rd ed.). New York: Macmillan.

Shepart, L.A. (1990). Inflated test score gains: Is the problem old norms or teaching the test? Educational Measurement: Issues & Practice, 9, 15-22.

Tittle, C.K. (1989). Validity: Whose construction is it in the teaching and learning context? Educational Measurement: Issues & Practice, 8, 5-13, 24.

Tittle, C.K., Kelly-Benjamin, K., & Sacks, J. (in press). The construction of validity and effects of large scale assessment in the schools. In R.E.Stake (ed)., Effects of changes in assessment policy. Vol 1, Advances in Program Evaluation. CT: JAI Press.

END

U.S. Dept. of Education

Office of Education
Research and
Improvement (OERI)

ERIC

Date Filmed

March 29, 1991